

IN THE CLAIMS

I claim:

1. A method of classifying Objects using a vector space having multiple preclassified data clusters comprising the steps of :
 - a. inputting a data stream that describes the Object;
 - b. abstracting the data stream to calculate an Object vector that characterizes the data stream;
 - c. identifying the data cluster, if any, within which the Object vector rests;
 - d. assigning to the Object the status of the identified data cluster or, if no cluster is identified, assigning to the Object the status of atypical.
2. The method of claim 1, wherein abstracting is performed by a process comprising selecting between 5 and 25 data points from the data stream.
3. The method of claim 1, wherein identifying is performed by a process comprising computing the Euclidean distance between the centroid of a data cluster and the Object vector.
4. The method of claim 1, wherein identifying is performed by a process comprising computing the normalized vector product of the Object vector and representing the centroid of a data cluster.
5. The method of claim 1, wherein each data cluster is preclassified as having one of two status conditions.
6. The method of claim 1, wherein each data cluster is preclassified as having one of three

status conditions.

7. The method of claim 1, wherein the data streams consist of between 1,000 and 20,000 data points.
8. The method of claim 1, wherein the length of the data streams consist of at least 1,000 data points.
9. A method of constructing a classifying algorithm by using a set of preclassified Objects, each Object being associated with a data stream, where the algorithm is characterized as having multiple data clusters of predetermined extent in a vector space of a fixed number of dimensions, comprising the steps of:
 - a. providing the set of the data streams associated with the preclassified Objects;
 - b. selecting an initial set of logical chromosomes that specify the location of a predetermine number of points of the data stream;
 - c. calculating an Object vector for each member of the set of data streams using each chromosome;
 - d. determining a fitness of each chromosome by finding the locations in the vector space of a multiplicity of non-overlapping data clusters of predetermined extent that maximize the number of Object vectors that rest in data clusters that contain only identically classified Object vectors, wherein the larger the number of such vectors the larger the fitness of the logical chromosome;
 - e. optimizing the set of logical chromosomes by an iterative process comprising reiteration of steps (c) and (d), terminating logical chromosomes with low fitness, replicating logical chromosomes of high fitness, recombination and random

modification of the chromosomes;

- f. terminating the iterative process and selecting a logical chromosome that allows for a optimally homogeneous set of non-overlapping data clusters, wherein the attributive status of each cluster of the optimally homogenous set is the classification of the Object vectors that rest within the data cluster; and
 - g. constructing a classifying algorithm that classifies an unknown Object by a process comprising calculating an unknown Object vector using the selected logical chromosome and classifying the unknown Object according to the attributive status of the data cluster of the optimally homogenous set of non-overlapping data clusters in which the unknown Object vector rests.
10. The method of claim 9, wherein the fixed number of dimensions is between 5 and 25.
 11. The method of claim 9, wherein the number of preclassified Objects is between 20 and 200.
 12. The method of claim 9, wherein the initial set of logical chromosomes is randomly selected.
 13. The method of claim 9, wherein the initial set of logical chromosomes consists of between 100 and 2,000 logical chromosomes.
 14. The method of claim 9, wherein the extent of each data cluster is equal.
 15. The method of claim 9, wherein the extent of each data cluster is determined by a Euclidean metric.
 16. The method of claim 15, wherein the extent of each data cluster in a dimension is a pre-determined fraction of the range of the Object vectors in the dimension.

17. The method of claim 9, wherein the metric that determines the extent of each data cluster is a function of a fuzzy AND match parameter with a vector characteristic of the data cluster.
18. The method of claim 9, wherein the location of each data cluster of the optimally homogenous set is the centroid of the Object vectors of preclassified Objects that rest in the data cluster.
19. The method of claim 9, wherein the location of each data cluster of the optimally homogenous set is the centroid of the Object vectors of preclassified Objects that rest in the data cluster.
20. The method of claim 9, wherein the location of each data cluster of the optimally homogenous set is the centroid of the Object vectors of preclassified Objects that rest in the data cluster.
21. A software product for a general purpose digital computer, accompanied by instructions that the product can be used to perform the method of claim 1 or of claim 9.
22. A software product, which performs or causes to be performed on a general purpose digital computer the method of claim 1 or claim 9.
23. A general purpose digital computer, programmed so as to performs or cause to be performed the method of claim 1 or claim 9.